

## Research on domain ontology in different granulations based on concept lattice

Xiangping Kang<sup>a,b</sup>, Deyu Li<sup>a,b,\*</sup>, Suge Wang<sup>a,b,c</sup>

<sup>a</sup> School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China

<sup>b</sup> Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Taiyuan, 030006 Shanxi, China

<sup>c</sup> School of Mathematics Science, Shanxi University, Taiyuan, 030006 Shanxi, China

### ARTICLE INFO

#### Article history:

Received 18 November 2010

Received in revised form 20 September 2011

Accepted 21 September 2011

Available online 29 September 2011

#### Keywords:

Concept lattice  
Granular computing  
Ontology building  
Ontology merging  
Ontology connection

### ABSTRACT

This paper introduces concept lattice and granular computing into ontology learning, and presents a unified research model for ontology building, ontology merging and ontology connection based on the domain ontology base in different granulations. In this model, as the knowledge in the lowest and most basic level, the domain ontology base is presented firstly, which provides a uniform technology for ontology learning on the whole; secondly, in order to better understand problems rather than be overwhelmed unnecessary details, granular computing is introduced to abstract and simplify domain ontology bases in complex domains. Moreover, the similarity of concepts in different granulations is introduced to help domain experts judging relations except for inheritance relation, and the similarity of ontologies in multi-granulations is introduced to measure the degree of connection of ontologies; finally, based on similarity models mentioned above, the ontology building, ontology merging and ontology connection can be obtained in different granulations with the help of domain experts. It is shown by instances that the application of the model presented in this paper is valid and practicable. Although there are still some problems in applications of this model (for example, ontology learning cannot dispense with the intervention of domain experts yet), this paper offers a new way for combining ontology learning and concept lattice.

Crown Copyright © 2011 Published by Elsevier B.V. All rights reserved.

## 1. Introduction

The term ontology is borrowed from philosophy. In 1993 Gruber proposed a most popular definition of ontology, that is, ontology is an explicit specification of a conceptual model [1]. Later Borst modified it slightly, and he presented that ontology is a formal and explicit specification of a shared conceptual model [2]. Although ontology is defined in different ways, researchers' recognition of it is unified from the perspective of essence. Since 1990s, research, exploitation and applications of ontology in computer area has become a hot issue, and ontology has gradually attract much attention of researchers in many areas, such as knowledge acquisition and representation, planning, process management, database framework integrated, natural language processing, business simulation etc. Along with going deeper into applications of ontology, some new mature and effective ontology learning methods are needed to support new demands, so it is necessary to make an active exploration on new methods. In recent years, many tools and techniques have been used to ontology research, and then a

mass of theories and methods have been achieved [3–11], such as the use of formal concept analysis cannot only weaken subjective effects in the process of ontology learning from developers, but also can automatically acquire implied concepts and relationships among concepts.

Concept consists of extent and intent, and based on this philosophy Professor Wille [12] proposed formal concept analysis (FCA), which is a method for finding, ordering and displaying concepts in early 1980s. Concept lattice is an effective tool in FCA, and it is very suitable for mining potential concepts of dates. It has been widely studied [14–20] and applied to machine learning [21], software engineering [22] and information retrieval [23].

Ontology aims to build a shared model for the objective world perceived by human, but concept lattice builds the model for artificial world rather than real world. We can build ontology for a certain domain when there is no data, but concept lattice must be built on a given data set. From the view of ISO704 standard shown in Fig. 1, it is not hard to see that the focus concerned by concept lattice and ontology is different. Concept lattice pays its concerns on the concept level, while ontology more concerns the presentation level. These two formal methods of ontology and concept lattice have little difference, but concept lattice viewed as a useful tool can be introduced into ontology. There are some common application fields of concept lattice and ontology in the philosophy,

\* Corresponding author at: School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China.

E-mail addresses: [kangxiangping@yahoo.cn](mailto:kangxiangping@yahoo.cn) (X. Kang), [lidyxu@yahoo.cn](mailto:lidyxu@yahoo.cn) (D. Li), [wsg@sxu.edu.cn](mailto:wsg@sxu.edu.cn) (S. Wang).

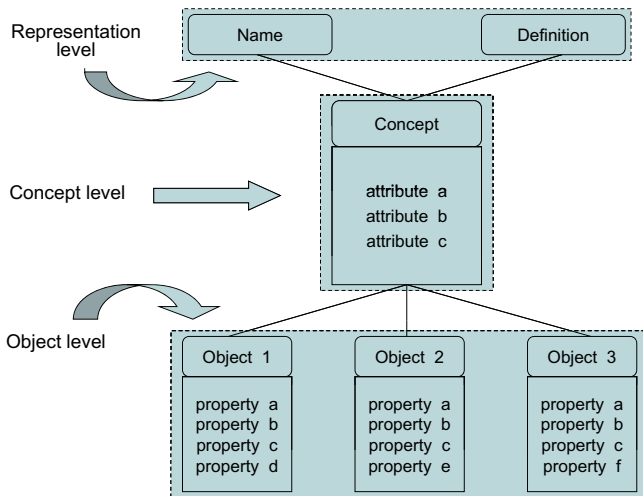


Fig. 1. The representation level, concept level and object level in ISO704.

information science, concept knowledge processing, knowledge presentation layer and so on, where concept lattice and ontology are complementary rather than competition or repulsion. In many applications concept lattice is the supplement to ontology. This paper introduces concept lattice into some key issues in the research of ontology, such as ontology building and ontology merging.

At present there are mainly two ways for ontology building: (1) ontology is described with the help of domain experts; (2) ontology is discovered from domain data. The first way of ontology building is completely performed manually. In some complex application areas this task will waste much time and energy, and it is subjective. Ontology built by different people even domain experts will differ in thousands ways, so the built ontology breaches the original intention of the introduction of ontology. The second way builds ontology by automatic or semiautomatic means, so the workload of building ontology manually can be reduced and the quality of ontology can be enhanced. In the second way, some important achievements of applying FCA to ontology building have been gained, the representative methods are as follows.

Cimiano et al. [24,25] analyzed the usage of words in a text by FCA to obtain the correspond background knowledge, then formed a domain ontology construction method. This method has following advantages: the “concept-attribute” relation is analyzed automatically from the domain text, which has great significance of reference; the synonym problem is solved by means of processing the same noun in context to be the same word. However, there also exist some limitations, for example: the result output from the language interpreter may not be always correct, which will bring mistakes to the domain ontology construction; relations of concepts are relatively single, which consider the inheritance relation only. But concepts interact by relations objectively, in the ontology construction other relations except the inheritance relation can be defined according to the need; it only aims at the pure text of the domain and do not consider the domain ontology construction based on other formal domain dates.

Tao [26] proposed a method of applying FCA to the ontology building. This method has following advantages: self-developing plugin can obtain a formal context from domain concepts and relationships automatically, and combined with the participation of domain experts the semiautomatic domain ontology construction is realized; redundant concepts in the taxonomic structure can be eliminated, and needed concepts can be obtained; the development idea proposed the domain ontology construction should depend on the feedback loop and should be improved continuously, which is worthy of reference. However, there also exist some

limitations, for example: based on the ontology modeling tool, the determinacy or limitation of modeling tool is brought into the domain ontology building; multi-valued attributes of the formal context are not fully considered so that the method is unsuitable for processing the multi-valued context; the plugin is taken as the media in the process of concepts conversing to the formal context, so that the complexity of conversion is increased. And some special relations in the initial model of ontology cannot be converted into corresponding formal contexts.

Haav [27,28] combined FCA with the rule-based language to perform a semi-automatic domain ontology building. This method has following advantages: a logic description of an ontology is realized, so that it is prone to reason and validate the ontology; an expansion mechanism of a domain ontology is proposed; non-taxonomic relations of a domain ontology and the ontology reasoning are fully considered. However, there also exist some limitations, for example: the conversion of initial ontology to set expressed by the first-order predicate logic needs FCA and rule language mapping that is complex, laborious and difficult to achieve; the formalization of domain ontology concepts is not enough to some extent; the concept extension is actually the number of domain texts, which results in the lexical gap of extension representation of domain ontology concepts. And the concept hierarchy relation established in this way is not conducive to express the relation between instances of the domain ontology.

Obitko et al. [29] applied FCA to the domain ontology building in GACR project. This method has following advantages: it provides a distributed ontology editing environment; a complete set of edit-modify mechanism for a formal context and a concept lattice is proposed, which has the value of reference; the visual concept lattice editing is realized. However, there also exist some limitations, for example: the whole process is such a continuous iteration process adjusted by adding or deleting concepts and attributes, so it is difficult to determine what contents need to be added or deleted and when the process is ended; the generation of formal context is required to be completed manually; it starts with empty objects and attributes, so the adding of objects and attributes is very complex, and the workload is heavy. Therefore, it is only appropriate for the small domain ontology construction. In general, the level of automatism of this method is lower than above methods.

The ontology merging is a process of merging two or more source ontologies into one target ontology. Manual ontology merging is time-consuming, laborious and easy to make mistakes by using traditional editing tools. Therefore, some scholars have proposed a number of new systems and frameworks to help knowledge engineers to merge ontologies, and they depend on heuristic methods matched with the grammar and semantic meaning adopted by ontology engineers in ontology merging. In these new methods, some important achievements of applying FCA to the ontology merging have been gained, the following is a brief introduction of these methods.

Stumme et al. [30] proposed the FCA-Merge method applying FCA to the ontology merging, which adopted bottom-up way to give the description of global process of ontology merging. Through the use of natural language processing technology, the FCA-Merge method can draw up instances from a text document in a specific domain, and then compute the corresponding context to create source ontology. In the context, objects are documents and attributes are concepts in the ontology, a concept and a document are related if the concept appears the document. Contexts can be connected after apposition, and then a concept lattice pruned can be computed by the TITANIC algorithm. Finally the generated concept lattice can be converted into a result ontology.

Ganter and Stumme [31] proposed the OntEx method based on FCA. Similar to FCA-merger, OntEx also gives measures to ensure that all possible merging can be considered. In practical

applications, OntEx needs to combine with a heuristic method, and interacts with knowledge engineers. But when the interaction amount is high, the cost of providing these measures is also huge.

Chen et al. [32] proposed a novel ontology merging method based on WordNet and FFCA techniques, called FFCA-Merge. According to the FFCA-Merge information, two extant ontologies with the same domain can be converted into a fuzzy ontology. The new fuzzy ontology is a single coherent ontology of high standards, and is more flexible than a general ontology.

In summary, combination methods of FCA and ontology research are different aiming to different application purposes. Therefore each method has its own advantages and disadvantages. Although there are various problems in practical applications, concept lattice offers a new solution to the problem of ontology.

By analyzing above ontology building and ontology merging methods carefully, a new FCA-based domain ontology research method proposed in this paper, which can be viewed as a necessary complement to existing methods mentioned above, it mainly solved problems as follows.

- Although FCA with a strong mathematical foundation can easily offer some implicated concepts from a given data set, a great number of concepts will be produced when the quantity of objects or attributes is larger. And further useful concepts may be overwhelmed. In order to avoid this case, granular computing is introduced to ontology research, which can overcome the impact of complex domains to some extent. That is, the larger granulation helps to hide some specific details, and then the problem can be treated from the overall picture.
- Normally, existing ontology learning models only adopt different logic methods locally, which cannot provide a uniform technology for ontology learning on the whole. To avoid the problem mentioned above, a unified ontology learning method based on domain ontology base is presented in this paper. As the knowledge in the lowest and most basic level, the domain ontology base not only offers a uniform technology for ontology learning on the whole, but also is convenient for knowledge sharing and reuse in the lowest level.
- Ontology merging is a process of merging two or more source ontologies into a new one, but normally people only are interested in the connection rather than merging between ontologies. In this case, the paper proposes the ontology connection in different granulations, and gives the corresponding measurement in multi-granulations, which measures the connection degree of ontologies.
- Although concept lattice is a hierarchy model and it can be used to realize the hierarchy model of ontology, relations in the hierarchy model of ontology are relatively single (namely, considering the inheritance relation only). This paper provides a new similarity technology, which can help experts judge relations except for inheritance relation.

The rest of the paper is organized as follows. Section 2 briefly recalls basic notions of domain ontology and concept lattice; Section 3 defines domain ontology base in different granulations with the help of domain experts; Section 4 describes the ontology building and ontology merging based on domain ontology case in different granulations; Section 5 discusses the ontology connection between different ontologies, and gives the corresponding measurement in multi-granulations, which reflects the connection degree; conclusions and the discussion of further work will close the paper in Section 6.

## 2. Basic notions of ontology and concept lattice

Domain ontology aims to capture relational domain knowledge, provides agreed understanding of domain knowledge, determines

recognized vocabulary and defines vocabulary and their relations explicitly at different levels of conceptualization. Human, database and application software share the domain knowledge by means of domain ontology.

A domain ontology is usually approximately defined as a binary group  $\mathcal{O} = (\mathcal{C}, \mathcal{R})$ , where  $\mathcal{C}$  is the set of all concepts,  $\mathcal{R}$  is the set of relations between concepts. Concepts and their relations constitute a directed graph. If  $T_i, T_j \in \mathcal{C}$  satisfying the relation  $\Delta \in \mathcal{R}$ , it is denoted by  $T_i \Delta T_j$ . For example, Fig. 2 is a fragment of a transportation ontology.

In the following we only provide the most basic notions of FCA, and more extensive introductions refer to [13].

In FCA, an elementary form of the representation of data is defined mathematically as formal context.

A formal context is a triple  $K = (G, M, I)$ , where  $G$  and  $M$  are sets, and  $I \subseteq G \times M$  is a binary relation. In the case, the members of  $G$  are called objects and the members of  $M$  are called attributes, and  $I$  is viewed as an incidence relation between objects and attributes. Accordingly, we write  $gIm$  or  $(g, m) \in I$  expressing “the object  $g$  has the attribute  $m$ ”.

Formal contexts are mostly represented by rectangular tables, rows of which are headed by object names and columns headed by attribute names. In the table, “ $\star$ ” means that the row object and the column attribute have some relation.

In  $K = (G, M, I)$ , for a set  $A \subseteq G$  of objects we define

$$A' = \{m \in M | gIm, \forall g \in A\}.$$

Correspondingly, for a set  $B \subseteq M$  of attributes we define

$$B' = \{g \in G | gIm, \forall m \in B\}.$$

(In the following, for  $g \in G$ , we write  $g'$  instead of  $\{g\}$ , and for  $m \in M$ , we write  $m'$  instead of  $\{m\}$ .)

$(A, B)$  is called a concept of  $K$ , if  $A' = B$  and  $B' = A$ . In this case,  $A$  is called the extent of  $(A, B)$ ,  $B$  is called the intent of  $(A, B)$ .  $\mathcal{B}(K)$  denotes the set of all concepts of  $K$ . If  $(A_1, B_1), (A_2, B_2) \in \mathcal{B}(K)$ , we define:

$$(A_1, B_1) \preceq (A_2, B_2) \iff A_1 \subseteq A_2 \iff B_1 \supseteq B_2,$$

where “ $\preceq$ ” is called the hierarchical order of concepts. The set of all concepts ordered in above way is called the concept lattice of  $K$ . In the absence of ambiguity, above concept lattice is still denoted as  $\mathcal{B}(K)$ .

Let  $(V, \preceq)$  be a concept lattice,  $Y \subseteq V$ , if every  $y \in V$  can be expressed as a supremum of a subset of  $Y$ , then,  $Y$  is called a supremum-dense of  $(V, \preceq)$ .

Let  $K = (G, M, I)$  be a formal context. For any object  $g \in G$ , we can define an object concept:

$$\gamma g = (g', g').$$

For any concept  $(A, B) \in \mathcal{B}(K)$ , it can be denoted by

$$(A, B) = \vee_{g \in A} \gamma g.$$

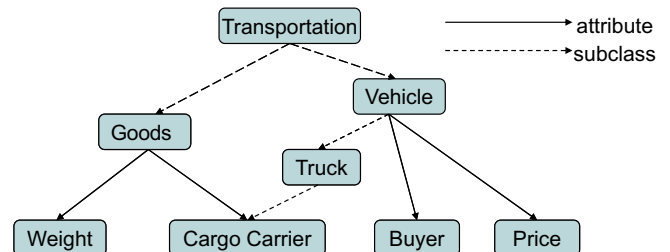


Fig. 2. A fragment of a transportation ontology.

Obviously, any  $(A, B) \in \mathcal{B}(K)$  can be denoted by the supremum of a certain subset of  $\gamma(G)$ , namely  $\gamma(G)$  is a supremum-dense in  $\mathcal{B}(K)$ .

### 3. The domain ontology base in different granulations

In the following, we can define a domain ontology base with the help of domain experts.

In  $\mathcal{O} = (\mathcal{C}, \mathcal{R})$ , let  $\Sigma \subseteq \mathcal{C}$ . If domain concept  $T$  can be obtained from  $\Sigma$  by the particular operation  $\tau$ , we say that  $T$  can be inferred from  $\Sigma$  by  $\tau$ . If any domain concept in  $\mathcal{O}$  can be inferred from  $\Sigma$  by  $\tau$ , we say that  $\Sigma$  is  $\tau$ -complete with respect to  $\mathcal{O}$ , where  $\tau$  expresses some mathematical operation of domain concepts.

**Domain ontology base:** in a domain  $\mathcal{O}$ , there exist a binary group  $\Sigma^{\mathcal{O}} = (\Sigma, R)$ , where  $\Sigma$  is a set of domain concepts and  $R$  is a fuzzy equivalence relation matrix on  $\Sigma$ . If  $\Sigma$  is  $\tau$ -complete, then  $\Sigma^{\mathcal{O}}$  is called an ontology base of the domain  $\mathcal{O}$ .

The  $\Sigma$  mentioned above is given by experts, that is, for any  $T \in \Sigma$ , domain experts give a specific and detailed description including name, attributes and other important information. In this paper, it only considers the name and attributes of  $T \in \Sigma$ , namely, experts define concepts of  $\Sigma$  by naming and assigning a set  $T^{\circ}$  of attributes for every concept  $T$ .

For example, if there is a given domain  $\mathcal{O}$ , then the corresponding domain ontology base defined by domain experts is shown in Table 1, the name of  $T \in \Sigma$  is simplified as  $i$  with  $i \in \{1 \cdots 8\}$ .

The similarity between any two domain concepts in  $\Sigma$  is defined as follows:

$$r_{ij} = \frac{|T_i^{\circ} \cap T_j^{\circ}|}{|T_i^{\circ} \cup T_j^{\circ}|},$$

where  $T_i, T_j \in \Sigma$ .

The fuzzy relation matrix is given as follows.

$$\tilde{R} = (r_{ij})_{\Sigma \times \Sigma},$$

where  $r_{ij}$  is defined above. If  $\tilde{R}$  satisfies

$$(1)r_{ii} = 1 \quad (2)r_{ij} = r_{ji} \quad (3)\bigvee_{k=1, \dots, |\Sigma|} (r_{ik} \wedge r_{kj}) \leq r_{ij},$$

we say  $\tilde{R}$  is a fuzzy equivalence relation matrix on  $\Sigma$ . Obviously  $\tilde{R}$  satisfies reflexivity and symmetry. If  $\tilde{R}$  does not satisfy transitivity, the fuzzy equivalence relation matrix of  $\Sigma$  can be obtained by the transitive closure algorithm. In the following, if  $\tilde{R}$  is a fuzzy equivalence relation matrix on  $\Sigma$ , then it is denoted by  $R$ .

The transitive closure is used to compute fuzzy equivalence matrix from a given fuzzy matrix. Given a fuzzy matrix  $\tilde{R}$ , its transitive closure  $R$  is computed as follows:

$$R = \tilde{R} \cup \tilde{R}^2 \cup \dots \cup \tilde{R}^{n-1}.$$

Based on above mentioned formula, the complexity to compute  $R$  by matrix multiplication as given in [37] is  $O(n^5)$ . Obviously the time complexity plays a major role in computing  $R$  from  $\tilde{R}$ . Hence for effectively descending the size of time complexity, various methods have been proposed to accelerate the computation. [38] et al. proposed an algorithm for computing the transitive closure of a fuzzy matrix which runs in time linear to the number of elements in the matrix, i.e.,  $O(n^2)$  if the matrix is an  $n \times n$  matrix.

**Table 1**  
A domain ontology base defined by domain experts.

Name	$T^{\circ}$	Name	$T^{\circ}$
1	a b d f h	5	b c e g h
2	b d f h i	6	b e g h j
3	c e h j	7	a b f i
4	a b d h i	8	b c g h j

The algorithm proposed is simple and easy to implement. In the following, we use the method in [38] to compute the transitive closure, and intermediate results will be omitted.

For example, referring to Table 1 we can obtain a fuzzy equivalence relation matrix on  $\Sigma$  by using the transitive closure algorithm, and the fuzzy equivalence relation matrix  $R$  is shown as follows

$$R = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix},$$

where  $n = 8$ , the specific experimental date is shown as follows

$$R = \begin{bmatrix} 1.00 & 0.67 & 0.25 & 0.67 & 0.25 & 0.25 & 0.50 & 0.25 \\ 0.67 & 1.00 & 0.25 & 0.67 & 0.25 & 0.25 & 0.50 & 0.25 \\ 0.25 & 0.25 & 1.00 & 0.25 & 0.50 & 0.50 & 0.25 & 0.50 \\ 0.67 & 0.67 & 0.25 & 1.00 & 0.25 & 0.25 & 0.50 & 0.25 \\ 0.25 & 0.25 & 0.50 & 0.25 & 1.00 & 0.67 & 0.25 & 0.67 \\ 0.25 & 0.25 & 0.50 & 0.25 & 0.67 & 1.00 & 0.25 & 0.67 \\ 0.50 & 0.50 & 0.25 & 0.50 & 0.25 & 0.25 & 1.00 & 0.25 \\ 0.25 & 0.25 & 0.50 & 0.25 & 0.67 & 0.67 & 0.25 & 1.00 \end{bmatrix}$$

By above discussion we can build a domain ontology base, in the following, an axiomatic definition of knowledge granulation in the domain ontology base is introduced.

Let  $\delta \in [0, 1]$ , the  $\delta$ -cut fuzzy equivalence relation matrix of  $R$  is shown as follows

$$R_{\delta} = (\tilde{r}_{ij}), \quad \text{where } \tilde{r}_{ij} = \begin{cases} 1 & r_{ij} \geq \delta, \\ 0 & r_{ij} < \delta. \end{cases}$$

It is not hard to see that  $R_{\delta}$  is still an equivalence relation matrix, obviously  $\Sigma/R_{\delta}$  is a partition of  $\Sigma$ .

As a powerful mechanism, Granular computing (GrC) was presented by Zadeh in 1996. He identified three basic concepts (namely, granulation, organization and causation) that underline the process of human cognition, where granulation involves decomposition of whole into parts, organization involves integration of parts into whole, and causation involves association of causes and effects. GrC is an umbrella term to cover any theories, methodologies, techniques, and tools that make use of granules in problem solving. As an effective tool for complex problem solving, it has potential applications in rough set theory, FCA, knowledge engineering, data mining, artificial intelligence, machine learning, etc., and has become an important research issue in information sciences field. More and more people have paid attentions to granular computing, and many pieces of nice work were accomplished in [39–43]. Although GrC in data mining has been explored widely and deeply by many scholars, there are few in the ontology learning. It is important to research approaches to ontology learning oriented GrC, which has theoretical significance and application prospect. The larger granulation helps to hide some specific details, and then the problem can be treated from the overall picture.

By referring [33], a definition of granulation is given as follows. Let

$$\Sigma/R_{\delta} = \{P_1, P_2, \dots, P_m\},$$

then the granulation of  $R_{\delta}$  is defined as

$$\rho(R_{\delta}) = \frac{1}{|\Sigma|^2} \cdot \sum_{i=1}^m |P_i|^2.$$

$P_i \in \Sigma/R_{\delta}$  is called granule. The  $\Sigma^{\mathcal{O}} = (\Sigma, R)$  in granulation  $\rho(R_{\delta})$  is denoted as  $\Sigma_{\delta}^{\mathcal{O}} = (\Sigma_{\delta}, R_{\delta})$ .

For example, in Table 1, the  $\Sigma^{\mathcal{O}}$  in different granulations are shown in Table 2. Notice that: any set  $\{p_1, p_2, \dots, p_n\}$  is abbreviated

to  $p_1 p_2 \dots p_n$  in following Tables. It is not hard to see the larger the value  $\delta$  is, the smaller the corresponding granulation is; conversely the smaller the value  $\delta$  is, the larger the corresponding granulation is.

Notice that, there is a necessary requirement objectively when the model in this paper is performed, that is the defined of domain ontology base needs the help of high-level domain experts to avoid the diffusion of incorrect information from lowest level.

**4. Domain ontology building and merging in different granulations**

Ontology is a formal and explicit specification of a shared conceptual model. In fact, most of formalization is not explicit but exist in documents, databases, or activities of brain quite vaguely. Therefore, for ontology building the difficulty is how to make implicit knowledge concrete and explicit in ontology, namely how to find out all possible abstract concepts and their relations.

As a branch of applied mathematics, concept lattice extracts all implicit concepts and their hierarchical order (also known as inheritance relation) from a given data automatically to form concept lattice. Concept consists of extent and intent, and concept lattice has explicit hierarchical order and plentiful semantic information. Ontology building usually starts with designing hierarchy model, which is the basis of all ontologies and ready for the domain ontology prototype model. Since concept lattice is a hierarchy model, concept lattice can be used to realize the hierarchy model of ontology approximately.

In the following, we will employ FCA to extract all implicit domain concepts and their inheritance relation from a given data automatically. If some domains are very complex, corresponding domain concepts will be great, some important domain concepts will be overwhelmed. Aiming to this case to some extent, we propose the research model based on domain ontology base in different granulations.

**Domain formal context:** in a domain  $\mathcal{O}$ ,  $\Sigma^\mathcal{O} = (\Sigma, R)$  is the domain ontology base,  $K_\mathcal{O} = (G_\mathcal{O}, M_\mathcal{O}, I_\mathcal{O})$  is the corresponding domain formal context, where elements of  $G_\mathcal{O}$  corresponds to names of concepts in  $\Sigma$  one by one.  $M_\mathcal{O}$  is the set of all attributes used by experts to characterize concepts of  $\Sigma$ .  $I \subseteq G_\mathcal{O} \times M_\mathcal{O}$  is viewed as an incidence relation between elements of  $G_\mathcal{O}$  and  $M_\mathcal{O}$ . For  $x \in G_\mathcal{O}$ ,  $m \in M_\mathcal{O}$ , we write  $(x, m) \in I_\mathcal{O}$  expressing “the concept  $x$  has the attribute  $m$ ”.

Normally, the description of a concept by extent and intent is redundant, because each of two parts determines the other. That is, if intents of two concepts are same, then it can be judged that two concepts are same. It is not hard to see that any intent of concept in  $\gamma(G_\mathcal{O})$  is equal to the  $T^p$  of concept  $T$  in  $\Sigma$  one by one, so object concepts in  $\gamma(G_\mathcal{O})$  can be viewed as domain concepts in  $\Sigma$  one by one (i.e.  $\gamma(G_\mathcal{O}) = \Sigma$ ). Since both of concepts in ontology and concepts in concept lattice root in philosophy, so they are same in essence, and corresponding concepts of concept lattice can be viewed as the quantification or specific definition of domain concepts in ontology. Obviously we can use  $\gamma(G_\mathcal{O})$  to replace  $\Sigma$ , and further use the operation  $\vee$  of  $\gamma(G_\mathcal{O})$  to simulate the operation  $\tau$  of  $\Sigma$ , where the operation  $\vee$  is defined as follows.

$$(A, B) = \bigvee_{g \in A} \gamma g, \quad \text{where } (A, B) \in \mathcal{B}(K_\mathcal{O}).$$

**Table 2**  
 $\Sigma^\mathcal{O}$  in different granulations.

$R_\delta$	$\Sigma_\delta$	$\rho(R_\delta)$
$R_{0.25}$	1 ~ 10	1.000
$R_{0.50}$	1247, 3568	0.500
$R_{0.67}$	124, 3, 568, 7	0.313
$R_{1.00}$	1, 2, 3, 4, 5, 6, 7, 8	0.125

In  $\mathcal{O} = (\mathcal{C}, \mathcal{R})$ , based on above discussions,  $\mathcal{B}(K_\mathcal{O})$  can be inferred from  $\gamma(G)$  by the operation  $\vee$ , in ideal cases (i.e.  $\Sigma$  defined by experts with no error), any concept in  $\mathcal{B}(K_\mathcal{O})$  named by domain experts is the domain concept in  $\mathcal{C}$ . That is, the  $\mathcal{C}$  is set of all concepts in  $\mathcal{B}(K_\mathcal{O})$  named by domain experts. So we can obtain  $\mathcal{C}$  from  $\gamma(G)$  with the help of domain experts. In the following,  $\gamma(G_\mathcal{O})$  is denoted as  $\Sigma$ .

For example, based on Table 1, the corresponding domain formal context is shown in Table 3, and then we can obtain the concept lattice shown in Fig. 3, where the name of domain concept in  $\Sigma$  is simplified to  $i$  with  $i \in \{1 \dots 8\}$ .

The quantization analysis of concepts in  $\mathcal{B}(K_\mathcal{O})$  are shown as follows.

**Concept base [13]:** in a domain  $\mathcal{O}$ , if  $(A, B) \in \mathcal{B}(K_\mathcal{O})$ , we say

$$(A, B)_\Sigma = \{\gamma g | g \in A\}$$

is the concept base of  $(A, B)$ .

If for all  $P \in \Sigma_\delta$ ,  $(A, B) \in \mathcal{B}(K_\mathcal{O})$  satisfies the following condition

$$1 - \frac{|P \cap (A, B)_\Sigma|}{|P|} \leq \theta,$$

where  $P \cap (A, B)_\Sigma \neq \emptyset$ , then we say that  $(A, B)$  is a concept in the granulation  $\rho(R_\delta)$ , the set of all concepts of  $\mathcal{B}(K_\mathcal{O})$  in the granulation  $\rho(R_\delta)$  is denoted as  $\mathcal{U}_\delta^\mathcal{O}$ . Correspondingly, the set of all granules concluded in  $(A, B)_\Sigma$  is denoted as  $(A, B)_{\Sigma(\delta, \theta)}$ .

Let  $\mathcal{U}_1$  and  $\mathcal{U}_2$  be two sets. If for any  $x \in \mathcal{U}_1$ , there always exists  $y \in \mathcal{U}_2$  such that  $x \subseteq y$ , we say that  $\mathcal{U}_2$  is coarser than  $\mathcal{U}_1$ , and it is denoted by  $\mathcal{U}_1 \leq \mathcal{U}_2$ .

From the above discussion, some conclusions can be inferred immediately as follows.

- If  $\theta_1 \leq \theta_2$ , then  $(A, B)_{\Sigma(\delta, \theta_2)} \leq (A, B)_{\Sigma(\delta, \theta_1)}$ .
- If  $\theta_1 \leq \theta_2$ , then  $\mathcal{U}_\delta^{\theta_2} \leq \mathcal{U}_\delta^{\theta_1}$ .
- When  $\delta \downarrow$  and  $\theta \downarrow$ , then  $|\mathcal{U}_\delta^\theta| \downarrow$ .
- When  $\delta \uparrow$  and  $\theta \uparrow$ , then  $|\mathcal{U}_\delta^\theta| \uparrow$ .

where “ $\downarrow$ ” means the value of parameter is becoming larger, “ $\uparrow$ ” means the value of parameter is becoming smaller. Users’ needs can be satisfied by adjusting values of parameters in the task of domain ontology learning, but  $\delta$  is viewed as the most important and basic parameter in this paper, others are only supplements to  $\delta$ .

In the following, we provide a new similarity model, which can help experts judge the relations except for inheritance relation. As the most basic relation, the inheritance relation constitutes backbone of the domain ontology together with domain concepts, other relations can be obtained based on the fact as follows.

- The similarity of concepts is contributed to judge relations except for inheritance relation in  $\mathcal{R}$  for domain experts.

For example, if the similarity between concepts  $T_i$  and  $T_j$  is 1, namely  $T_i$  and  $T_j$  are same in essence, then they must satisfy synonymy relation; if the similarity between  $T_i$  and  $T_j$  is 0, obviously  $T_i$

**Table 3**  
A domain formal context.

	$a$	$b$	$c$	$d$	$e$	$f$	$g$	$h$	$i$	$j$
1	★	★		★		★		★		
2		★		★		★		★	★	
3			★		★			★		★
4	★	★		★				★	★	
5		★	★		★		★	★		
6		★			★		★	★		★
7	★	★				★			★	
8		★	★				★	★		★

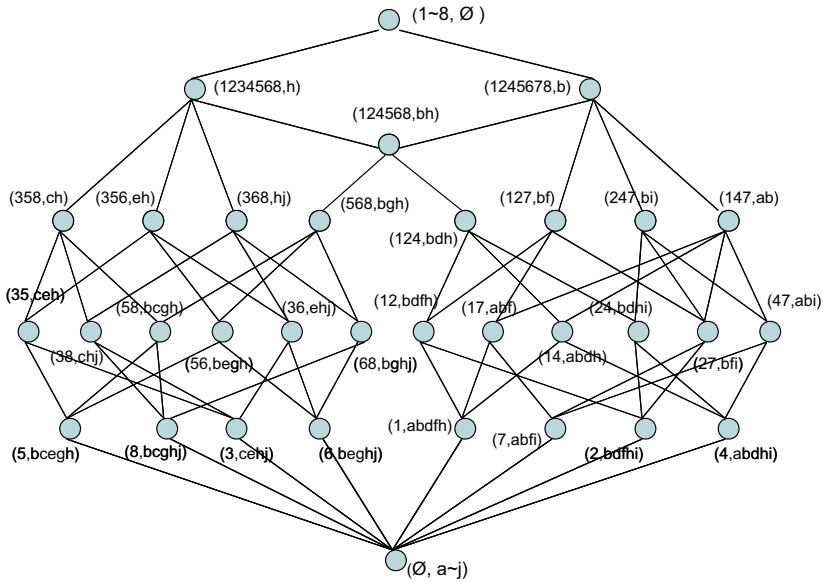


Fig. 3. The concept lattice based on Table 3.

and  $T_j$  must not satisfy the synonymy relation and inheritance relation; if the similarity between  $T_i$  and  $T_j$  is 0.8, and the relation between  $T_i$  and  $T_j$  may be inheritance relation or whole-part relation, then the probability of the former is greater than the later in most cases, and so on.

At present there are many similarity measure models which can be generally classified as two classes: they are sequential metric space model and set theory matching model [34] respectively. An example of the former is Shepard model that is based on probability distribution. The latter can be further classified as Geometric model, Transformational model, Featur model and so on. Geometric model calculates the similarity between concepts by computing the distance between feature vectors of entity in n-dimensional space. Transformational model calculates the similarity between concepts by the number of needed steps of one entity converting to another. Featur model calculates the similarity between concepts by considering the number of sets of entity's common features. A typical example of featur model is the Tversky model [35]

$$sim(a, b) = \frac{f(A \cap B)}{\bar{f}(A \cap B) + \alpha f(A - B) + \beta f(B - A)},$$

where  $A$  and  $B$  are sets of features implicated in the concept  $a$  and  $b$ ,  $f$  denotes a measure function of the set of features. Based on Tversky model Rodrigue and Egenhofer proposed a similarity measure model of concepts by common character and difference character [36]

$$sim(a, b) = \frac{|A \cap B|}{|A \cap B| + \alpha(a, b)|A - B| + (1 - \alpha(a, b))|B - A|},$$

where  $A$  and  $B$  are feature sets implicated in the concept  $a$  and  $b$  with  $0 \leq \alpha(a, b) \leq 0.5$ . This paper presents a new similarity model, that is,  $\theta$ -similarity in the granulation  $\rho(R_\delta)$ , which relies on granules contained in the concept base. Essentially this model still belongs to feature model.

Let  $(A_1, B_1), (A_2, B_2) \in \mathcal{W}_\delta^\theta$ , the  $\theta$ -similarity in the granulation  $\rho(R_\delta)$  between  $(A_1, B_1)$  and  $(A_2, B_2)$  is defined as

$$Sim_{(\delta, \theta)}((A_1, B_1), (A_2, B_2)) = \frac{|\bigcap_{i=1,2} (A_i, B_i)_{\Sigma(\delta, \theta)}|}{|\bigcup_{i=1,2} (A_i, B_i)_{\Sigma(\delta, \theta)}|}.$$

Obviously, based on the above formula the following conclusion can be obtained easily.

- Let  $(A_i, B_i) \in \mathcal{W}_\delta^\theta$  with  $i = 1, 2, 3$ , if  $(A_1, B_1) \leq (A_2, B_2) \leq (A_3, B_3)$ , then  $Sim_{(\delta, \theta)}((A_1, B_1), (A_3, B_3)) \leq Sim_{(\delta, \theta)}((A_1, B_1), (A_2, B_2))$ .

From the above conclusion we can see the  $\theta$ -similarity in the granulation  $\rho(R_\delta)$  mentioned above is feasible. For example, let us start by evaluating the similarity of two sibling concepts in the Fig. 3, where  $(\delta, \theta) = (0.67, 0.34)$ . That is,  $(124568, bh)$  and  $(368, hj)$ . In this case, we have

$$Sim_{(\delta, \theta)}((124568, bh), (368, hj)) = \frac{|\{124, 568\} \cap \{3, 568\}|}{|\{124, 568\} \cup \{3, 568\}|} = \frac{1}{3} = 0.33.$$

Similarity increases if we consider a concept and one of its direct descendent. In fact, consider again  $(124568, bh)$ , and let us evaluate the similarity with the child  $(568, bgh)$ . The following holds

$$Sim_{(\delta, \theta)}((124568, bh), (568, bgh)) = \frac{|\{124, 568\} \cap \{568\}|}{|\{124, 568\} \cup \{568\}|} = \frac{1}{2} = 0.50.$$

Obviously, similarity decreases in the case of concepts that are not directly related. For instance, consider one of the previous concepts, that is  $(568, bgh)$  and its ancestor  $(1234568, h)$ . We have

$$Sim_{(\delta, \theta)}((568, bgh), (1234568, h)) = \frac{|\{568\} \cap \{124, 3, 568\}|}{|\{568\} \cup \{124, 3, 568\}|} = \frac{1}{3} = 0.33.$$

Now, based on the  $\theta$ -similarity in the granulation  $\rho(R_\delta)$ ,  $\mathcal{W}_\delta^\theta$  and the hierarchical order of concepts in  $\mathcal{W}_\delta^\theta$ , we can obtain the conceptual hierarchy model of  $\mathcal{W}_\delta^\theta$  automatically, which includes two parts as follows:

- Concepts in  $\mathcal{W}_\delta^\theta$ .
- The hierarchical order of concepts in  $\mathcal{W}_\delta^\theta$ .
- The similarity of concepts in  $\mathcal{W}_\delta^\theta$  based on the  $\theta$ -similarity in the granulation  $\rho(R_\delta)$ .

For example, in Fig. 3, when  $\delta = 0.67$  and  $\theta = 0.34$ , the corresponding conceptual hierarchy model of  $\mathcal{W}_\delta^\theta$  is shown in Fig. 4. For simplifying Fig. 4, we only computes the similarity of

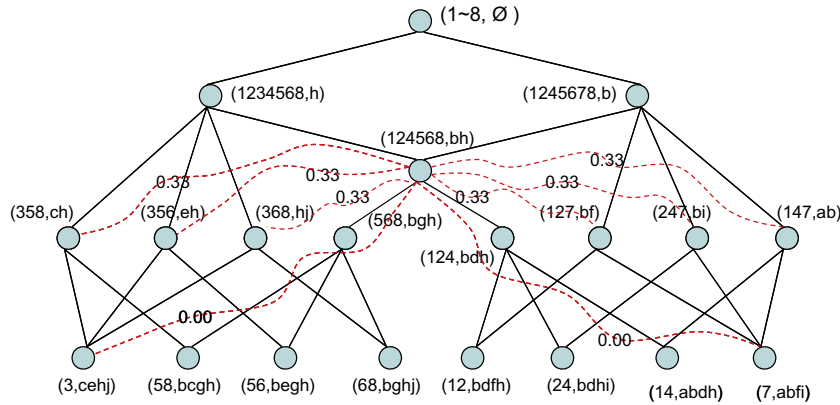


Fig. 4. The conceptual hierarchy model of  $\mathcal{U}_\delta^0$  with  $\delta = 0.67$  and  $\theta = 0.34$ .

(124568,bh) with others as an example. In the same way, when  $\delta = 0.50$  and  $\theta = 0.25$ , the conceptual hierarchy model of  $\mathcal{U}_\delta^0$  is shown in Fig. 5.

In a domain  $\mathcal{O}$ ,  $\mathcal{R}$  can be defined as various types such as synonymy relation, whole-part relation, inheritance relation and so on. But in the conceptual hierarchy model of  $\mathcal{U}_\delta^0$ , we only obtain the inheritance relation between concepts. Based on the conceptual hierarchy model of  $\mathcal{U}_\delta^0$  and opinions of domain experts, the domain ontology building in the granulation  $\rho(R_\delta)$  is shown as follows.

**Algorithm** (The domain ontology building in the granulation  $\rho(R_\delta)$ )

Input:  $\Sigma, \delta, \theta$  and opinions of domain experts;

Output:  $\mathcal{O}_\delta = (\mathcal{C}_\delta, \mathcal{R}_\delta)$ ;

Steps are shown as follows

1. Building the domain ontology base  $\Sigma^e = (\Sigma, R)$ , based on which  $K_e$  is obtained and corresponding  $\mathcal{B}(K_e)$  is built;
2. Based on  $\Sigma_\delta^e$ , opinions of experts are introduced to name concepts in  $\mathcal{U}_\delta^0$ , namely generating the set  $\mathcal{C}_\delta$  of domain concepts;
3. The hierarchical order of concepts is viewed as the inheritance relation, by calculating the similarity between concepts in  $\mathcal{U}_\delta^0$  based on the  $\theta$ -similarity in the granulation  $\rho(R_\delta)$ , other relations can be judged by domain experts. That is,  $\mathcal{R}_\delta$  is generated;
4. Output  $\mathcal{O}_\delta = (\mathcal{C}_\delta, \mathcal{R}_\delta)$ ;
- 5.End.

Ontology merging means to merge several existent domain ontologies to eliminate repetitive and incongruous parts. Different ontologies are merged into a ontology with more rational concept system and stronger knowledge representation ability. It's meaningless to merge ontology from domains without cross, so ontologies to be merged must come from two cross domains.

For the sake of describing problems conveniently, some formal symbols in this paper are defined as follows: let  $\mathcal{O}^\alpha$  and  $\mathcal{O}^\beta$  be two different domain ontologies,  $(\Sigma^\alpha, R^\alpha)$  and  $(\Sigma^\beta, R^\beta)$  be corresponding ontology bases. The connection between ontologies will generate a new domain ontology  $\mathcal{O} = (\mathcal{C}, \mathcal{R})$ , the corresponding ontology base is  $\Sigma^e = (\Sigma, \mathcal{R})$  and domain formal context is  $K_e$ , where  $\Sigma = \Sigma^\alpha \cup \Sigma^\beta$ .

Let  $\mathcal{O}^\alpha$  and  $\mathcal{O}^\beta$  be two different domain ontology, a new domain ontology is generated by merging  $\mathcal{O}^\alpha$  and  $\mathcal{O}^\beta$  based on  $\Sigma = \Sigma^\alpha \cup \Sigma^\beta$  by using the algorithm of domain ontology building in the granulation  $\rho(R_\delta)$ , the process of merging is not discussed in detail.

For example, there is another domain, the corresponding domain ontology base is shown in Table 4. if the domain determined by Table 1 is denoted as  $\mathcal{O}^\alpha$  and the domain determined by Table 4 is denoted as  $\mathcal{O}^\beta$ , then based on  $\Sigma = \Sigma^\alpha \cup \Sigma^\beta$ , we can obtain a new domain ontology  $\mathcal{O}$  after merging, some computing results as follows.

1. The corresponding domain formal context  $K_e$  is shown in Table 5.
2. Corresponding conceptual hierarchy models of  $\mathcal{U}_\delta^0$  in  $\mathcal{O}$  are shown in Fig. 6 ( $\delta = 0.67$  and  $\theta = 0.00$ ) and Fig. 7 ( $\delta = 0.43$  and  $\theta = 0.00$ ).

There are two types storage structures of data in computer and linear structure is one of them. In the domain of linear structure, the corresponding domain ontology base is noted as  $\mathcal{O}^\alpha$  in Table 6. By using the above ontology building algorithm, the “Data’s linear structure” ontologies in different granulations are shown in Fig. 8 ( $\delta = 0.75$  and  $\theta = 0.40$ ) and Fig. 9 ( $\delta = 0.40$  and  $\theta = 0.20$ ) separately. The other storage structure of data is non-linear structure, the corresponding domain ontology base is noted as  $\mathcal{O}^\beta$  in Table 6. Then merging the  $\mathcal{O}^\alpha$  and the  $\mathcal{O}^\beta$  generates a new domain ontology “Data’s logical structure”. The “Data’s logical structure” ontologies are shown in Fig. 10 ( $\delta = 0.75$  and  $\theta = 0.40$ ) and Fig. 11 ( $\delta = 0.15$  and  $\theta = 0.43$ ). In Table 6, the symbols are interpreted as follows:  $a_1$  = linear,  $a_2$  = sequential,  $a_3$  = double,  $a_4$  = stack,  $a_5$  = string,  $a_6$  = special,  $a_7$  = queue,  $a_8$  = single,  $a_9$  = linked,  $a_{10}$  = dynamic,  $a_{11}$  = circular,  $a_{12}$  = common,  $a_{13}$  = static,  $a_{14}$  = non-linear,  $a_{15}$  = tree-like,  $a_{16}$  = graphic and  $a_{17}$  = set.

**5. The ontology connection in the granulation  $\rho(R_\delta)$**

In the following, we will introduce ontology connection, which is an progress of connecting two or more source ontologies by some new concepts in  $\mathcal{U}_\delta^0$ . That is, in  $\mathcal{U}_\delta^0$ , we will find some key concepts, which have high correlations with two domain ontologies. If some domains are very complex, the connection between ontologies are very complicated, corresponding concepts will be very great, some important key concepts will be overwhelmed. In order to avoid this case, we propose the  $(\theta, \vartheta)$ -ontology connection in the granulation  $\rho(R_\delta)$ , namely which can overcome the impact of the complex domain ontologies to some extent.

**$\vartheta$ -condition:** in  $\Sigma = \Sigma^\alpha \cup \Sigma^\beta$ , let  $0 \leq \vartheta \leq 1$ , if  $(A, B) \in \mathcal{U}_\delta^0$  satisfies following conditions

$$1 - \frac{|(A, B)_{\Sigma(\delta, \theta)} \cap \Sigma_\delta^\alpha|}{|(A, B)_{\Sigma(\delta, \theta)}} \leq \vartheta \quad \text{and} \quad 1 - \frac{|(A, B)_{\Sigma(\delta, \theta)} \cap \Sigma_\delta^\beta|}{|(A, B)_{\Sigma(\delta, \theta)}} \leq \vartheta,$$

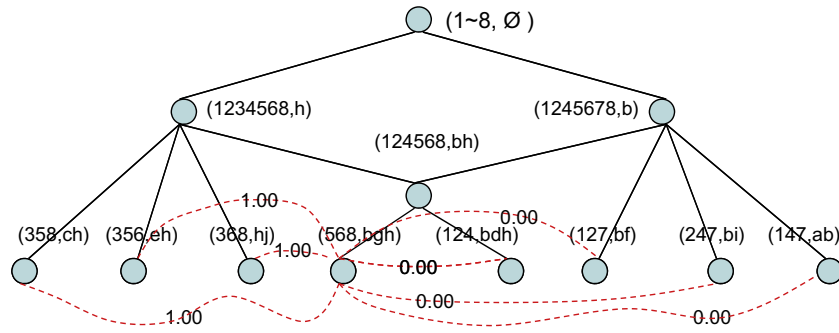


Fig. 5. The conceptual hierarchy model of  $\mathcal{W}_\delta^\theta$  with  $\delta = 0.50$  and  $\theta = 0.25$ .

Table 4

A domain ontology base defined by domain experts.

Name	$T^a$	Name	$T^b$
1	afkln	7	afin
2	fikln	9	aeikln
3	eklm	10	aelmn
4	aikln	11	aklmn

Table 5

A domain formal context.

	a	b	c	d	e	f	g	h	i	j	k	l	m	n
1	*	*		*		*		*			*	*		*
2		*		*		*		*	*		*	*		*
3			*	*				*		*	*	*	*	*
4	*	*		*				*	*	*	*	*		*
5		*	*	*			*	*	*		*			*
6		*		*			*	*	*	*	*			*
7	*	*		*		*		*	*	*	*			*
8		*	*	*			*	*	*	*	*			*
9	*			*							*	*	*	*
10	*			*							*	*	*	*
11	*			*							*	*	*	*

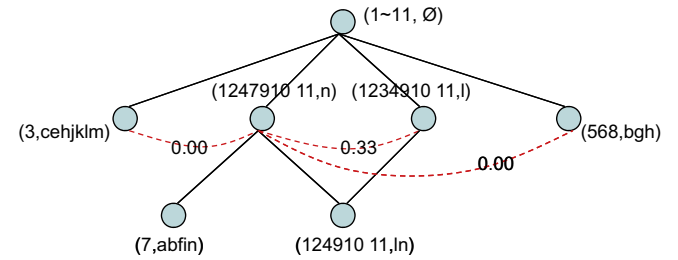


Fig. 7. The conceptual hierarchy model of  $\mathcal{W}_\delta^\theta$  with  $\delta = 0.43$  and  $\theta = 0.00$ .

is denoted as  $\mathcal{W}_\delta^{\theta,\vartheta} = \{(A, B) \in \mathcal{W}_\delta^\theta \mid (A, B) - \vartheta\}$ . Then the following conclusion holds.

- If  $\vartheta_1 \leq \vartheta_2$ , then  $\mathcal{W}_\delta^{\theta,\vartheta_2} \subseteq \mathcal{W}_\delta^{\theta,\vartheta_1}$ .
- When  $\delta \downarrow$ ,  $\theta \downarrow$  and  $\vartheta \downarrow$ , then  $|\mathcal{W}_\delta^{\theta,\vartheta}| \downarrow$ .
- When  $\delta \uparrow$ ,  $\theta \uparrow$  and  $\vartheta \uparrow$ , then  $|\mathcal{W}_\delta^{\theta,\vartheta}| \uparrow$ .

The similarity between different domain ontologies reflects the degree of connection between two ontologies to some extent. If the similarity is 0, they have no connection; if the similarity is 1, they are the same ontology. The similarity between different domain ontologies can guide to set the parameter  $\vartheta$  rationally in the process of  $(\theta, \vartheta)$ -ontology connection in the granulation  $\rho(R_\delta)$ . For example, if the similarity between ontologies is small, we should set  $\vartheta$  larger, conversely we should set  $\vartheta$  smaller. In this way some unimportant concepts can be removed, so that some important key concepts satisfying  $\vartheta$ -condition avoid being overwhelmed.

Let  $\mathcal{O}^\alpha$  and  $\mathcal{O}^\beta$  be two different domain ontology,  $\Sigma = \Sigma^\alpha \cup \Sigma^\beta$  satisfies the condition that for any  $\delta_i$  and  $\delta_j$  ( $1 \leq i, j \leq n$ ),  $\Sigma_{\delta_i} \neq \Sigma_{\delta_j}$  holds, where  $n$  is the maximum value satisfying above condition. Then the similarity between  $\mathcal{O}^\alpha$  and  $\mathcal{O}^\beta$  in multi-granulations is defined as

$$Sim(\mathcal{O}^\alpha, \mathcal{O}^\beta) = \frac{1}{n} \cdot \sum_{k=1}^n (S_k / N_k),$$

where  $S_k = |\{P \in \Sigma_{\delta_k} \mid \exists Q \in \Sigma_{\delta_k}^\beta, P \cap Q \neq \emptyset \text{ and } P \cap S \neq \emptyset\}|$ ,  $N_k = |\Sigma_{\delta_k}|$ .  $N_k$  denotes the number of granules contained in  $\Sigma_{\delta_k}$ .  $S_k$  denotes the number of some granules satisfying the condition that one part elements are from  $\Sigma^\alpha$ , and the other part elements are from  $\Sigma^\beta$ .  $Sim(\mathcal{O}^\alpha, \mathcal{O}^\beta)$  can avoid the possible large deviation existing in the single-granulation efficiently.

For example, there are  $\Sigma_\delta^\alpha$ ,  $\Sigma_\delta^\beta$  and  $\Sigma_\delta$  ( $\Sigma = \Sigma^\alpha \cup \Sigma^\beta$ ) in Table 7, then  $Sim(\mathcal{O}^\alpha, \mathcal{O}^\beta)$  is equal to

$$\frac{1}{4} \times \left\{ \frac{|\{2, 6\}|}{|\{1, 2, 3, 4, 5, 6, 7, 8\}|} + \frac{|\{12, 46\}|}{|\{12, 35, 46, 78\}|} + \frac{|\{1246\}|}{|\{1246, 3578\}|} + \frac{|\{1 \sim 8\}|}{|\{1 \sim 8\}|} \right\} = \frac{9}{16}.$$

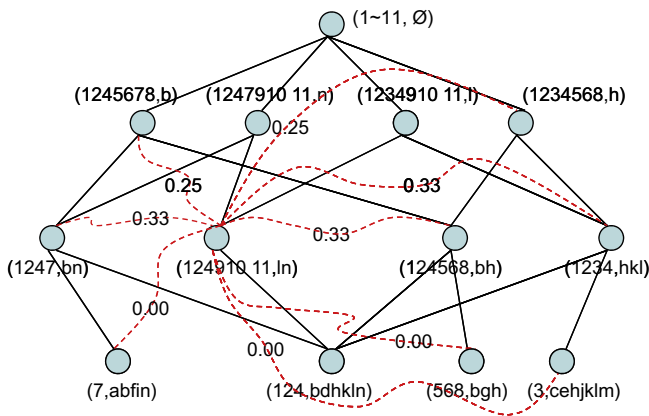


Fig. 6. The conceptual hierarchy model of  $\mathcal{W}_\delta^\theta$  with  $\delta = 0.67$  and  $\theta = 0.00$ .

we say that  $(A, B)$  satisfies  $\vartheta$ -condition, and it is denoted by  $(A, B) - \vartheta$ . Namely, if  $(A, B) - \vartheta$ , then  $(A, B)_{\Sigma(\delta, \vartheta)}$  consists of two parts

1. One part is from  $\Sigma_\delta^\alpha$ .
2. The other part is from  $\Sigma_\delta^\beta$ .

and proportion of each part is larger than  $1 - \vartheta$ . If  $(A, B) - \vartheta$ , it reflects that the concept  $(A, B)$  has high correlations with two domains to a certain extent. The set of all concepts satisfying the  $\vartheta$ -condition



**Table 6**  
A domain ontology base defined by domain experts.

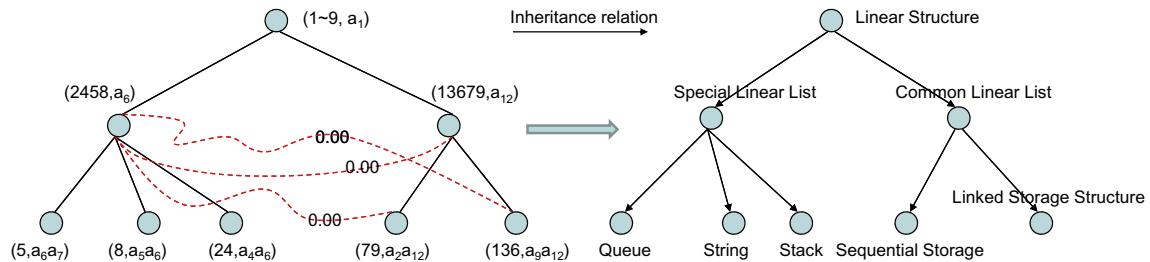
The domain ontology base of $\mathcal{O}^\alpha$			The domain ontology base of $\mathcal{O}^\beta$		
Name	$T^\alpha$		Name	$T^\beta$	
1	Single Link List	$a_1, a_8, a_9, a_{12}$	6	Circular Link List	$a_1, a_9, a_{11}, a_{12}$
2	Sequential Stack	$a_1, a_2, a_4, a_6$	7	Dynamic Allocation	$a_1, a_2, a_{10}, a_{12}$
3	Double Link List	$a_1, a_3, a_9, a_{12}$	8	String	$a_1, a_5, a_6$
4	Link Stack	$a_1, a_4, a_6, a_9$	9	Static Allocation	$a_1, a_2, a_{12}, a_{13}$
5	Queue	$a_1, a_6, a_7$			
10	Tree	$a_{14}, a_{15}$	12	Set	$a_{14}, a_{17}$
11	Graph	$a_{14}, a_{16}$			

**Table 7**  
The domain ontology base in different granulations.

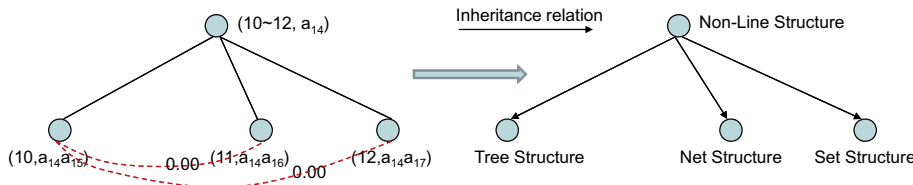
$\delta$	$\Sigma_\delta^\alpha$	$\Sigma_\delta^\beta$	$\Sigma_\delta$
$\delta_1$	1, 2, 4, 6	2, 3, 5, 6, 7, 8	1, 2, 3, 4, 5, 6, 7, 8
$\delta_2$	12, 46	25, 36, 78	12, 46, 35, 78
$\delta_3$	1246	2578, 36	1246, 3578
$\delta_4$	1246	235678	1 ~ 8

Concepts in  $\mathcal{U}_\delta^{\theta, \vartheta}$  are named by experts, then experts determine the relations between concepts by referring the similarity base on  $\theta$ -similarity in the granulation  $\rho(R_\delta)$ , where “relations of concepts” means

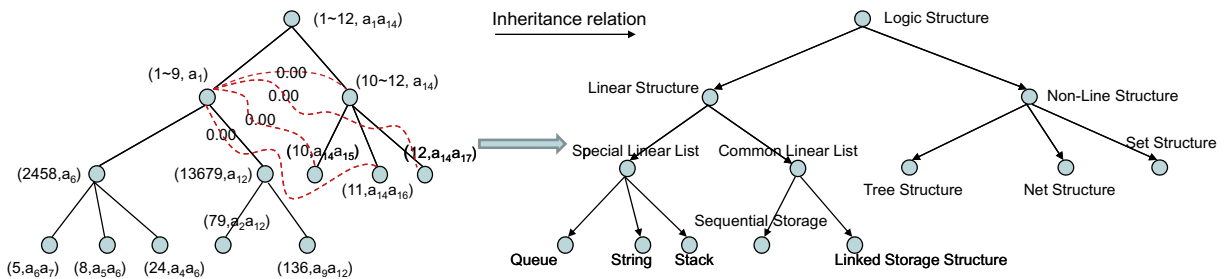
1. Relations between concepts in  $\mathcal{O}^\alpha$  and concepts in  $\mathcal{U}_\delta^{\theta, \vartheta}$ .
2. Relations between concepts in  $\mathcal{O}^\beta$  and concepts in  $\mathcal{U}_\delta^{\theta, \vartheta}$ .
3. Relations of concepts in  $\mathcal{U}_\delta^{\theta, \vartheta}$ .



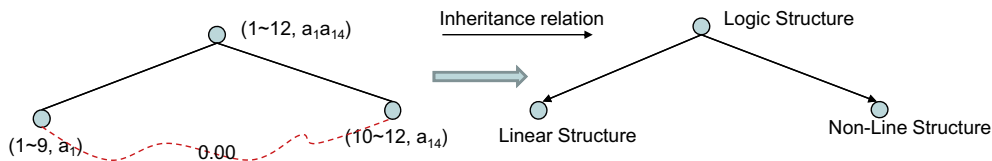
**Fig. 8.** The “Data’s linear structure” with  $\delta = 0.75$  and  $\theta = 0.40$ .



**Fig. 9.** The “Data’s linear structure” with  $\delta = 0.40$  and  $\theta = 0.20$ .



**Fig. 10.** The “Data’s logical structure” with  $\delta = 0.75$  and  $\theta = 0.40$ .



**Fig. 11.** The “Data’s logical structure” with  $\delta = 0.15$  and  $\theta = 0.43$ .

Now, based above knowledge, the ontology connection can be obtained. the corresponding process of ontology connection is not discussed in detail.

## 6. Conclusions

This paper introduces concept lattice and GrC into ontology research, and presents a unified research model for ontology building, ontology merging and ontology connection based on domain ontology base in different granulations, and provides a detailed description of this overall process. In this model, it mainly obtains conclusions as follows: (1) A unified domain ontology base is provided for ontology research. As the knowledge in the lowest and most basic level, the domain ontology base not only offers a uniform technology for ontology learning on the whole, but also is convenient for knowledge sharing and reuse in the lowest level; (2) GrC is introduced to ontology research, which can overcome the impact on the application of FCA caused by the time complexity and space complexity problem to some extent, it helps to find useful information and avoids users getting lost in the complex information; (3) A new similarity between concepts is given in different granulations, which can help experts judge relations except for inheritance relation; (4) The connection between different ontologies is proposed, and the corresponding similarity in multi-granulations to measure the degree of connection of ontologies. Although the FCA-based domain ontology learning proposed in this paper is only a starting point and a lot of subsequent study is needed, but it offers a new way or guideline for ontology learning. How to combine concept lattice with domain ontology more rationally and reduce human judgement is one focus of our research in the future.

## Acknowledgements

The work is supported by the National Natural Science Foundation of China (60970014, 61070100 and 60875040), the Natural Science Foundation of Shanxi, China (2010011021-1) and the Graduate Innovation Project of Shanxi Province, China (20103004).

## References

- [1] T. Gruber, Translation approach to portable ontology specifications, *Knowledge Acquisition* 5 (1993) 199–220.
- [2] W.N. Borst, Construction of engineering ontologies for knowledge sharing and reuse. Ph.D. Thesis, University of Twente, Enschede, 1997.
- [3] C.S. Lee, M.H. Wang, Ontology-based computational intelligent multi-agent and its application to CMMI assessment, *Applied Intelligence* 30 (3) (2009) 203–219.
- [4] C.S. Lee, M.H. Wang, J.J. Chen, Ontology-based intelligent decision support agent for CMMI project monitoring and control, *International Journal of Approximate Reasoning* 48 (1) (2008) 62–76.
- [5] S. Calegari, F. Farina, Fuzzy ontologies and scale-free networks analysis, *International Journal of Computer Science & Applications* 4 (2) (2007) 125–144.
- [6] H. Alani, S. Kim, D. Millard, M. Weal, W. Hall, P. Lewis, N. Shadbolt, Automatic ontology-based knowledge extraction from web documents, *IEEE Intelligent Systems* 18 (1) (2003) 14–21.
- [7] R. Gacitua, P. Sawyer, P. Rayson, A flexible framework to experiment with ontology learning techniques, *Knowledge-Based Systems* 21 (3) (2008) 192–199.
- [8] T.T. Quan, S.C. Hui, A.C.M. Fong, Automatic fuzzy ontology generation for semantic help-desk support, *IEEE Transactions on Industrial Informatics* 2 (3) (2006) 155–164.
- [9] R.C. Chen, J.Y. Liang, R.H. Pan, Using recursive ART network to construction domain ontology based on term frequency and inverse document frequency, *Expert Systems with Applications* 34 (1) (2008) 488–501.
- [10] J. Köhler, S. Philippi, M. Specht, A. Rüegg, Ontology based text indexing and querying for the semantic web, *Knowledge-Based Systems* 19 (8) (2006) 744–754.
- [11] A. Formica, Ontology-based concept similarity in formal concept analysis, *Information Sciences* 176 (18) (2006) 2624–2641.
- [12] R. Wille, Restructuring lattice theory: an approach based on hierarchies of concepts. in: I. Rival (ed.), *Ordered Sets*, Dordrecht, Reidel, 1982, pp. 445–470.
- [13] B. Ganter, R. Wille, *Formal Concept Analysis: Mathematical Foundations*, Springer-Verlag, Berlin, 1999.
- [14] L. Wei, J.J. Qi, Relation between concept lattice reduct and rough set reduct, *Knowledge-Based Systems* 23 (8) (2010) 934–938.
- [15] L. Wang, X. Liu, A new model of evaluating concept similarity, *Knowledge-Based Systems* 21 (8) (2008) 842–846.
- [16] A. Burusco, R. Fuentes-Gonzalez, Construction of the L-fuzzy concept lattice, *Fuzzy Sets and Systems* 97 (1) (1998) 109–114.
- [17] J.S. Mi, Y. Leung, W.Z. Wu, Approaches to attribute reduct in concept lattices induced by axialities, *Knowledge-Based Systems* 23 (6) (2010) 504–511.
- [18] J. Medina, M. Ojeda-Aciego, J. Ruiz-Calviño, Formal concept analysis via multi-adjoint concept lattices, *Fuzzy Sets and Systems* 160 (2) (2009) 130–144.
- [19] L. Li, J. Zhang, Attribute reduction in fuzzy concept lattices based on the T implication, *Knowledge-Based Systems* 23 (6) (2010) 497–503.
- [20] X. Wang, W.X. Zhang, Relations of attribute reduction between object and property oriented concept lattices, *Knowledge-Based Systems* 21 (5) (2008) 398–403.
- [21] B. Zupa, M. Bohance, Learning by discovering concept hierarchies, *Artificial Intelligence* 109 (1–2) (1999) 211–242.
- [22] U. Dekel, Revealing Java class structure with concept lattices. Master's Thesis. Technion-Israel Institute of Technology, 2003.
- [23] P. Valtchev, R. Missaoui, R. Godin, M. Meridji, Generating frequent itemsets incrementally: two novel approaches based on Galois lattice theory, *Journal of Experimental and Theoretical Artificial Intelligence* 14 (2–3) (2002) 115–142.
- [24] P. Cimiano, G. Stumme, A. Hotho, J. Tane, Conceptual knowledge processing with formal concept analysis and ontologies, in: *Proceedings of the Second International Conference on Formal Concept Analysis (ICFCA 04)*, Springer, 2004, pp. 189–207.
- [25] P. Cimiano, S. Staab, J. Tane, Automatic acquisition of taxonomies from text: FCA meets NLP, in: *Proceedings of the International Workshop on Adaptive Text Extraction and Mining*, 2003, pp. 10–17.
- [26] G. Tao, Using formal concept analysis for ontology structuring and building, ICIS, Nanyang Technological University, 2003.
- [27] H.M. Haav, A semi-automatic method to ontology design by using FCA, in: V. Snasel, R. Belohlavek (Eds.), *Concept Lattices and their Applications*, Proceedings of the 2nd International CLA Workshop, TU of Ostrava, 2004, pp. 13–25.
- [28] H.M. Haav, An application of inductive concept analysis to construction of domain-specific ontologies, in: B. Thalheim, Gunar Fiedler (Eds.), *Emerging Database Research in East Europe, Proceedings of the Pre-conference Workshop of VLDB 2003*, Computer Science Reports, Brandenburg University of Technology at Cottbus, 14(3) (2003) 63–67.
- [29] M. Obitko, V. snasel, J. Smid, Ontology Design with Formal Concept Analysis, Edited by Vaclav Snasel, Radim Belohlavek. in: *Proceedings of the CLA 2004 International Workshop on Concept Lattices and their Applications*, Ostrava, Czech Republic, Sep. 2004, pp. 111–119.
- [30] G. Stumme, M. Adche, A FCA-Merge: bottom-up merging of ontologies. in: *Proceedings of the Seventeenth International Conference on Artificial Intelligence (IJCAI'01)*, Seattle, USA, 2001, pp. 225–230.
- [31] B. Ganter, G. Stumme, Creation and merging of ontology top-levels, in: A.de Moor, W. Lex, B. Ganter (Eds.), *Conceptual Structures for Knowledge Creation and Communication. Proc. ICCS'03, LNAI, 2746*, Springer, Heidelberg, 2003, pp. 131–145.
- [32] R.C. Chen, C.T. Bau, C.J. Yeh, Merging domain ontologies based on the WordNet system and fuzzy formal concept analysis techniques, *Applied Soft Computing* 11 (2) (2011) 1908–1923.
- [33] J.Y. Liang, Z.Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 12 (1) (2004) 37–46.
- [34] J.B. Tenenbaum, T.L. Griffiths, Generalization, similarity, and Bayesian inference, *Behavioral and Brain Sciences* 24 (2001) 629–640.
- [35] A. Tversky, Features of similarity, *Psychological Review* 84 (4) (1977) 327–352.
- [36] M.A. Rodriguez, M.J. Egenhofer, Determining semantic similarity among entity classes from different ontologies, *IEEE Transactions on Knowledge and Data Engineering* 15 (2003) 442–456.
- [37] L.A. Zadeh, Fuzzy sets, *Information and Control* 8 (1965) 338–353.
- [38] H.S. Lee, An optimal algorithm for computing the max-min transitive closure of a fuzzy similarity matrix, *Fuzzy Sets and Systems* 123 (2001) 129–136.
- [39] D. Dubois, H. Prade, On possibility theory, formal concept analysis and granulation, *Survey. Appl. Comput. Math.* 10 (1) (2011) 10–19.
- [40] T.Y. Lin, Granular computing on binary relations I: data mining and neighborhood systems, II: rough sets representations and belief functions, in: L. Polkowski, A. Skowron (Eds.), *Rough Sets in Knowledge Discovery*, 1, Physica-Verlag, Heidelberg, 1998, pp. 107–140.
- [41] Y.Y. Yao, Granular computing: basic issues and possible solutions, in: *Proceedings of the 5th Joint Conference on Information Sciences*, 2000, pp. 186–189.
- [42] S.B. Roh, S.K. Oh, W. Pedrycz, A fuzzy ensemble of parallel polynomial neural networks with information granules formed by fuzzy clustering, *Knowledge-Based Systems* 23 (3) (2010) 202–219.
- [43] J. Zhao, L. Liu, Construction of concept granule based on rough set and representation of knowledge-based complex system, *Knowledge-Based Systems* 24 (6) (2011) 809–815.